



Briefing

Online-Diffamierung

Stichworte: Hate Speech, kriminelle Diffamierung, Verleumdung, üble Nachrede, Belästigung, Beleidigung, Blasphemie, Majestätsbeleidigung

Thematik und Begriffe: ein Überblick

Unter Diffamierung zu verstehen ist die Verbreitung einer falschen Aussage, die den Ruf einer Person, eines Unternehmens, eines Produkts, einer Gruppe, einer Regierung, einer Religion oder einer ganzen Nation schädigt. Diffamierungen werden je nach nationaler Gesetzgebung zivil- und/oder strafrechtlich geregelt. Diffamierende Kommentare sind oft rassistisch oder sexistisch und richten sich an bestimmte Personen oder Gruppen. Online-Diffamierung ist daher ein Oberbegriff für das Phänomen der gruppenbezogenen Menschenfeindlichkeit oder Aufruhr im Internet und in den sozialen Medien.

Marginalisierte Gruppen sind bevorzugte Objekte von Diffamierung. Aber auch etablierte Gruppen oder Einzelpersonen können zum Ziel von Verleumdungskampagnen werden.

Diffamierungen sind Zuschreibungen durch Dritte. Die Akteure selbst betrachten ihre Aussagen als wahr oder legitim. In diesem Themenkomplex geht es also weitgehend um das soziale Verständnis dessen, was als Verleumdung wahrgenommen wird und als solche bekämpft werden soll. Der Online-"Pranger" fungiert als erweiterte Form der Öffentlichkeit. Dies reicht von privaten Personengruppen in sozialen Netzwerken über global tätige Influencer bis hin zu NGOs und staatlichen Akteuren.

Sie alle profitieren von:

- Echtzeit-Kommunikation über beliebige Distanzen
- Anonymität der (Zahl der) Akteure
- grenzüberschreitenden Reichweiten
- Multiagentensystemen
- Zielgruppen-Kommunikation (Targeting)

Online-Diffamierung wird verstärkt als gesellschaftliches Problem erkannt, dem die Politik gemeinsam mit zivilgesellschaftlichen Akteuren begegnen muss. Rassistisch-sexistische Hasskampagnen nationalistischer Kreise sind derzeit die offensichtlichste Form der organisierten Online-Diffamierung - nicht nur in diesem Land. Als Teil der vierten Gewalt werden Journalist*innen zum Teil doppelt aufgrund von Geschlecht und Beruf diskriminiert und einem Kreis von diffamierenden Strategien ausgesetzt (siehe Abbildung).



Das Internet als Medium der Globalisierung ist zum Schauplatz und Träger populistischer anti-globalistischer Strategien geworden. Die Beobachtung solcher Tendenzen in verschiedenen Ländern innerhalb und außerhalb Europas und die grenzüberschreitende Vernetzung relevanter Aktionen haben dieses Thema auf die Agenda der globalen Kooperationsforschung gesetzt.

Ohne Anspruch auf Vollständigkeit zu erheben, werden in diesem *Briefing* Aspekte zusammengefasst, die für die öffentliche Diskussion relevant sein können.

Online-Belästigung von Journalistinnen



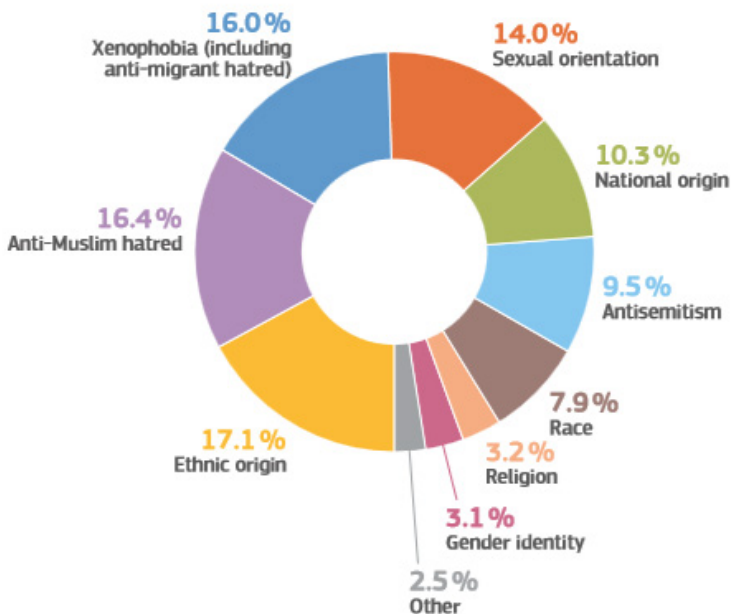
Die Herausforderung in Europa und Deutschland

EU: der 'Code of conduct' zu Bekämpfung illegaler *Hate Speech* online

Im Rahmen des EU-Internetforums vereinbarte die EU-Kommission im Mai 2016 mit den führenden IT-Unternehmen Facebook, Microsoft, Twitter und YouTube einen Verhaltenskodex zur Bekämpfung illegaler Hassreden im Internet. Auf freiwilliger Basis sollen die Internetunternehmen dafür Sorge tragen, Meldungen von Nutzern innerhalb von 24 Stunden zu bearbeiten, die Nutzer besser dabei unterstützen, diffamierende Inhalte im Netz zu melden und über die Regeln in der Online-Community aufklären.

- Der Vereinbarung traten 2018 des weiteren Instagram, Google+, Snapchat und Dailymotion bei.
- Die Vereinbarung wird regelmäßig evaluiert. Die Ergebnisse der letzten Evaluierung vom Januar 2018 sind:
 - Durchschnittlich 70% der gemeldeten Hassreden (insgesamt 2.982 Meldungen) wurden von den IT-Unternehmen gelöscht.
 - 81% der gemeldeten Hassbeiträgen konnten innerhalb von 24 Stunden gesichtet werden.
 - Die Berichterstattung über Hate Speech wurde vereinfacht und die Berichtsverfahren transparenter gestaltet.
 - Die Zusammenarbeit mit zivilgesellschaftlichen Organisationen zur Sensibilisierung für Hate Speech im Internet wurde intensiviert.
 - Hate Speech bezieht sich meist auf rassistische Hassreden gegen ethnische Minderheiten, Migrant*innen und Geflüchtete. Die häufigsten Gründe für die Meldung von Hate Speech waren ethnische Herkunft (17,1%), anti-muslimischer Hass (16,4%) und Fremdenfeindlichkeit (16%).

Zielscheibe Minderheiten: EU Code of Conduct third monitoring (2018): Benachrichtigungen nach Grund der Hassrede (in %)



Zielscheibe: Jugendliche

Eine aktuelle Umfrage vor der Landtagswahl in Hessen bestätigt Befunde anderer Studien: Jugendliche, hier die Altersgruppe der 18-24-Jährigen, sind von Hasskommentaren und Bedrohungen in Online-Netzwerken besonders betroffen. 42% berichteten von emotionalem Stress und sexueller Belästigung. Über 50% der 18-bis 44-Jährigen, das entspricht drei Alterskohorten, berichteten, dass sie einmal oder mehrfach beleidigt wurden. Die Androhung körperlicher Gewalt haben 40% der Jüngeren bereits mindestens einmal erlebt.

Zielscheibe: Journalist*innen

Eine Studie des Instituts für interdisziplinäre Konflikt- und Gewaltforschung zur "Wahrnehmung von und Erfahrungen mit Angriffen unter Journalist*innen" berichtet, dass insgesamt 42% der befragten Journalist*innen 2016 hasserfüllte Angriffe aus ihrem Publikum erlebten. Mehr als ein Fünftel der Befragten (26%) berichten, dass sie "mehrmals" bis "regelmäßig" angegriffen wurden. Die Studie bezieht sich auch auf die polizeiliche Kriminalstatistik des Bundesinnenministeriums aus dem Jahr 2015. Diese zeigt, dass kriminelle Hasskommentare im Internet gegenüber dem Vorjahr um 176% zugenommen haben. Eine Umfrage unter 66 Tageszeitungsvollredaktionen ergab, dass 27 von ihnen bestimmte Inhalte nicht mehr auf Facebook veröffentlichen, mehr als die Hälfte der befragten Zeitungsredaktionen gaben an, mit der Moderation ihrer Online-Foren überfordert zu sein.

AfD: Studie findet, dass soziale Medien als Verbreitungsmechanismus zwischen Hate Speech und realer Gewalt fungieren können

Die Studie von Müller/Schwarz untersuchte den Zusammenhang zwischen Facebook-Likes und dem Facebook-Account der *Alternative für Deutschland*. Hassdelikte nahmen in Zeiten, in welchen der AfD-Account viele Beiträge gegen Geflüchtete veröffentlichte, zu. Die Autoren entdeckten außerdem eine Korrelation zwischen Donald Trumps Tweets und der Anzahl von Hassverbrechen gegen Muslime in US-Bundesstaaten mit hoher Twitter-Nutzung.

Umstritten: das deutsche Netzwerkdurchsetzungsgesetz (NetzDG)

Um soziale Netzwerke im Internet zur Reaktion auf rechtswidrige Inhalte zu zwingen, trat am 1. Januar 2018 das sogenannte Netzwerkdurchsetzungsgesetz (NetzDG) in Kraft.

Das Gesetz ist umstritten. Viele Kritiker fürchten, dass die Betreiber der Netzwerke aufgrund der kurzen 24-Stunden-Frist vorsorglich zu viel löschen und damit die Meinungsfreiheit der Nutzer einschränken. Wird ein Beitrag oder ein Konto unberechtigt gelöscht, haben die Betroffenen kaum eine Chance, gegen die Netzwerke vorzugehen. Widerspruchsmöglichkeiten sieht das NetzDG nicht vor. Satire oder Ironie werden oft nicht erkannt (Gefahr des Overblocking). Hinzukommt, dass die Richtlinienkompetenz durch das NetzDG in die Hand der Konzerne gelegt wird.

Globale Trends zur Meinungsfreiheit und Medienentwicklung

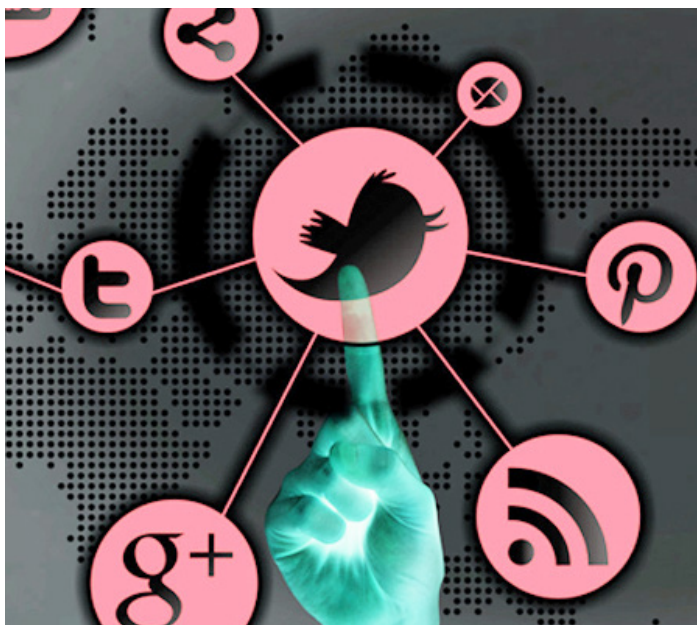
Zwei Arten von Beziehungen zwischen globalen digitalen Plattformen und nationaler Regulierung werden beobachtet:

- Die meisten Staaten der Welt haben das Wachstum globaler Großunternehmen ermöglicht, die über den Rahmen staatlicher Regulierung hinausgehen und bei denen Formen von Selbstregulierung in Absprache mit den Regierungen im Trend liegen.
- In manchen Teilen der Welt wird die nationalstaatliche Souveränität bekräftigt und versucht, die Plattformen im Einklang mit strengen Kontrollen der lokalen etablierten Medien zu "domestizieren". Als Teil dieses letztgenannten Trends sind Internet-Provider im Besitz der Regierung oder werden kontrolliert oder befinden sich in den Händen von regierungsnahen Unternehmen, und die Datenlokalisierung ist vorgeschrieben.

Diffamierung: zwei Verständnisse

Viele Länder haben Diffamierung als Teil eines Trends zur Öffnung der Gesellschaft und mehr Meinungsfreiheit entkriminalisiert. In den letzten Jahren ist die Online-Entwicklung jedoch einem gegenteiligen Trend gewichen. Laut einem UNESCO-Bericht haben Länder in allen Weltregionen begonnen, Diffamierungen zunehmend zu kriminalisieren, indem sie ihre jeweilige Gesetzgebung auf Online-Inhalte ausdehnen. Auch Cyberkriminalität und Anti-Terrorismus-Gesetze wurden weltweit verabschiedet.

Die Unabhängigkeit der Medien steht aufgrund der komplexen Verbindungen zwischen politischer Macht und Regulierungsbehörden zunehmend unter Druck. Gleichzeitig fungieren Algorithmen im Netz außerhalb transparenter Governance-Vereinbarungen als proprietäre Lösungen zur Kontrolle von Online-Inhalten. Dies kann eine Herausforderung für die Legitimität der im globalen Kontext vereinbarten Verfahren darstellen.



Online-Diffamierung von Rohingyas

Im März teilte eine Untersuchungsmission der Vereinten Nationen in Myanmar mit, dass Soziale Medien "wesentlich zum Ausmaß der Aggressivität" in der breiten Öffentlichkeit gegen Rohingya-Muslime beigetragen habe und dass das Phänomen *Online-Hate Speech* "sicherlich ein Teil davon" sei. Yanghee Lee, Sonderberichterstatterin für die Situation der Menschenrechte in Myanmar, erklärte: "Wir wissen, dass die ultra-nationalistischen Buddhisten ihre eigenen Facebook-Accounts haben und darüber wirklich viel Gewalt und Hass gegen die Rohingya oder andere ethnische Minderheiten schüren." Reuters untersuchte die anhaltenden Schwierigkeiten der Plattformen bei der Überwachung ihrer Inhalte und ihre Zurückhaltung, das Problem angemessen anzugehen.

Afrika

Auf dem afrikanischen Kontinent haben mindestens vier Staaten zwischen 2012 und 2017 Diffamierung entkriminalisiert.

Der Afrikanische Gerichtshof für Menschenrechte und Völkerrecht stellte in *Kanaté gegen Burkina Faso* fest, dass Diffamierungsgesetze ein "Überbleibsel des Kolonialismus" seien.

"Niemand kann für wahre Aussagen haftbar gemacht werden". (Afrikanische Kommission für Menschenrechte und Rechte der Völker).

Falsche, nicht-objektive Informationen

"Allgemeine Verbote für die Verbreitung von Informationen, die auf vagen und zweideutigen Ideen beruhen, einschließlich falscher Nachrichten, sogenannter "Fake News" oder "nicht objektiver Informationen", sind mit den internationalen Normen für Beschränkungen der Meinungsfreiheit unvereinbar (...) und sollten abgeschafft werden."

(UN-Sonderberichterstatter für freie Meinungsäußerung, OSZE, OAS und Sonderberichterstatter für Meinungsfreiheit der Afrikanischen Kommission für Menschenrechte und Rechte der Völker, 2017.)

Eine Analyse arabischer Gesetze

Duffy/Alkazemi analysierten Bestimmungen in Ägypten, Jordanien, Kuwait, Libanon, Libyen und den Vereinigten Arabischen Emiraten. Diffamierung wird hier in der Regel mit einer Geldstrafe oder einer Gefängnisstrafe von ein bis zwei Jahren bestraft. Das Cyberkriminalitätsgesetz der Vereinigten Arabischen Emirate hingegen sieht drei bis fünfzehn Jahre Gefängnis und "harte Geldstrafen" für die Veröffentlichung von Dingen vor, die den Ruf einer langen Liste von Führungskräften beschädigen könnten.

Wahr aber diffamierend

Viele Gesetzgeber betrachten Wahrheit als Mittel gegen Diffamierungsbeschuldigungen. Im libanesischen Strafrecht, Artikel 583, wird jedoch den Personen, die der Verleumdung beschuldigt werden, die Möglichkeit verweigert, Informationen damit zu rechtfertigen, dass sie wahr seien. Das Gesetz wurde geschrieben, um wahrheitsgemäße Aussagen mit diffamierender Wirkung zu unterbinden (Duffy/Alkazemi).

Kommentare und der Abschreckungseffekt – eine Herausforderung

Zwischen 2013 und 2016 haben 20% der Online-Publikationsplattformen Kommentare entweder abgeschaltet oder offline gestellt, nachdem selbst zeitaufwändige Maßnahmen nicht funktioniert haben (World Editors Forum). Zahlreiche Nutzer*innen berichten von ähnlichen Erfahrungen mit den sozialen Medien: Die Menge an toxischem Material fühlt sich unerträglich an, und Menschen, die sich in diffamierenden Umgebungen bewegen und exponieren, sind emotionalen und physischen Gefahren ausgesetzt.

Der *Guardian* untersuchte die 70 Millionen Kommentare, die zwischen 1999 und 2016 auf seiner Website erfasst wurden. Von diesen Kommentaren wurden etwa zwei Prozent wegen missbräuchlichen oder störenden Verhaltens gesperrt.

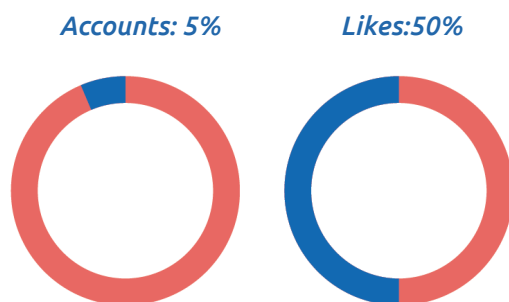
Bemerkenswerterweise sind unter den zehn Journalisten und Journalistinnen, welche das höchste Maß an Missbrauch und "abschätzigem Trolling" erfuhren, acht Frauen und zwei dunkelhäutige Männer.

Es gibt aber auch Erfolge: insbesondere im Bereich der Problemanalyse und Kommunikationsstrategien.

Eine quantitative Analyse der Likes für hasserfüllte Beiträge in den sozialen Medien ergab, dass 5% der Konten 50% der Likes für hasserfüllte Beiträge im Datensatz hervorbrachten (dargestellt im Diagramm unten). Dies zeigt: Zahlen sind nicht immer das, wonach sie aussehen. Auf diese Weise wird eine Mehrheit suggeriert und Legitimität proklamiert; unter denen, die so handeln, sind auch staatliche Akteure.

Freiwillige beteiligen sich an Online-Diskussionen, um gefälschte Nachrichten oder hasserfüllte Inhalte aufzudecken. In Deutschland etwa agiert die Gruppe #wirsindhier.

Weltweit versuchen die Verleger, qualitativ hochwertige Kommentare ihrer Nutzer*innen zu fördern. Laut einer Umfrage aus dem Jahr 2016 (World Editors Forum) begrüßte eine Vielzahl der Verleger (22%) eine verstärkte Moderation, um die Leser nicht zu vergraulen. Nur 8% haben die Kommentarspalten vollständig geschlossen.



Unternehmen und Plattformen gegen Hate Speech

Hoaxmap - Neues aus der Gerüchteküche

Plattform in Deutschland, die versucht, falsche Meldungen aufzudecken: <https://hoaxmap.org>. Auf einer interaktiven Karte findet man hier zahlreiche Falschmeldungen, die örtlich zugeordnet und widerlegt werden.

No Hate Speech Movement Deutschland

<https://no-hate-speech.de>

Task Force "Umgang mit rechtswidrigen Hassbotschaften im Internet"

<https://www.fair-im-netz.de>

Canada

Canadian Anti-Hate Network

<https://www.antihate.ca>

United States

Southern Poverty Law Center: <https://www.splcenter.org>

Color of Change: <https://colorofchange.org>

Quellen

European Commission: Countering illegal hate speech online http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=54300.

Siegert, S., 2016: Exklusive Journalist-Umfrage. Nahezu jede zweite Zeitungsredaktion schränkt Online Kommentare ein. In: *journalist online*, 01.03.2016.

Institut für Demokratie und Zivilgesellschaft (IDZ) im Auftrag von Campact (2018). #Hass im Netz: Der schleichende Angriff auf unsere Demokratie. Eine repräsentative Untersuchung in Hessen.

Müller, Karsten and Schwarz, Carlo, Fanning the Flames of Hate: Social Media and Hate Crime (May 21, 2018). <http://dx.doi.org/10.2139/ssrn.3082972>.

Müller, Karsten and Schwarz, Carlo, Making America Hate Again? Twitter and Hate Crime Under Trump (March 30, 2018). <http://dx.doi.org/10.2139/ssrn.3149103>.

Netzwerkdurchsetzungsgesetz (NetzDG) and Critique (in German): <https://www.bpb.de/dialog/netzdebatte>.

UNESCO. 2018. World Trends in Freedom of Expression and Media Development: 2017/2018 Global Report, Paris.

Matt J. Duffy & Mariam Alkazemi (2017). Arab Defamation Laws: A Comparative Analysis of Libel and Slander in the Middle East, *Communication Law and Policy*, 22:2, 189-211, DOI: 10.1080/10811680.2017.1290984.

Online Defamation of Rohingyas <https://www.bbc.com/news/technology-43385677>; <https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/>.

Philip Kreißel, Julia Ebner, Alexander Urban, Jakob Guhl (2018). Hass auf Knopfdruck. Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz. Hrsg. Institute for Strategic Dialogue (ISD) und #ichbinhier. https://www.isdglobal.org/wp-content/uploads/2018/07/ISD_Ich_Bin_Hier_2.pdf.

Code of Conduct on countering illegal hate speech online. Results of the 3rd monitoring exercise. January 2018.

World Editors Forum (2016). Do Comments Matter? Global Online Commenting Study 2016. Frankfurt.



Centre for

**Global
Cooperation
Research**

Käte Hamburger Kolleg / Centre for Global Cooperation Research (KHK/GCR21)

A Central Research Institute of the University of Duisburg-Essen

Director: Sigrid Quack

Co-Directors: Tobias Deibel, Dirk Messner, Jan Aart Scholte

Schifferstraße 44, 47059 Duisburg

Phone: +49 (0)203 379 5230 Fax: +49 (0)203 379 5276

E-Mail: info@gcr21.uni-due.de